

MACHINA DIXIT

ULISES CORTÉS

*La inteligencia
en esta nueva aldea
es de artificio¹.*

Siguendo la línea trazada por Hobbes², hace más de ciento sesenta años Samuel Butler escribió, en plena revolución industrial, una visionaria carta al editor del diario *The Press*³, de la localidad neozelandesa de Christchurch, en la que anunciaba que los humanos estábamos en la senda de crear una especie *sucesora*, en la línea de la evolución que anunció Darwin, y, además, previó que las *máquinas* adquirirían la supremacía de éstas sobre el mundo y sus habitantes.

También propone una *guerra a muerte* contra las máquinas, instando a la destrucción de toda la maquinaria para salvar a la humanidad de una eventual subyugación, aunque, en 1860, reconoce que ya es *demasiado* tarde.

La especulación de Butler sobre la posibilidad de que los humanos se conviertan para las máquinas en lo que los animales domésticos son para los humanos es un argumento central y matizado en su carta *Darwin entre las máquinas*. Butler establece una analogía directa entre la domesticación de los animales y la evolución de la relación entre los humanos y las máquinas cada vez más sofisticadas.



Esta visión resulta especialmente relevante hoy, cuando las tecnologías basadas en la inteligencia artificial — como los dispositivos digitales, las redes sociales o los modelos masivos de lenguaje como los de las familias ChatGPT⁴ o Llama— ya configuran entornos donde las personas quedan, en muchos sentidos, atrapadas o condicionadas por sistemas algorítmicos que median sus decisiones, relaciones y percepciones.

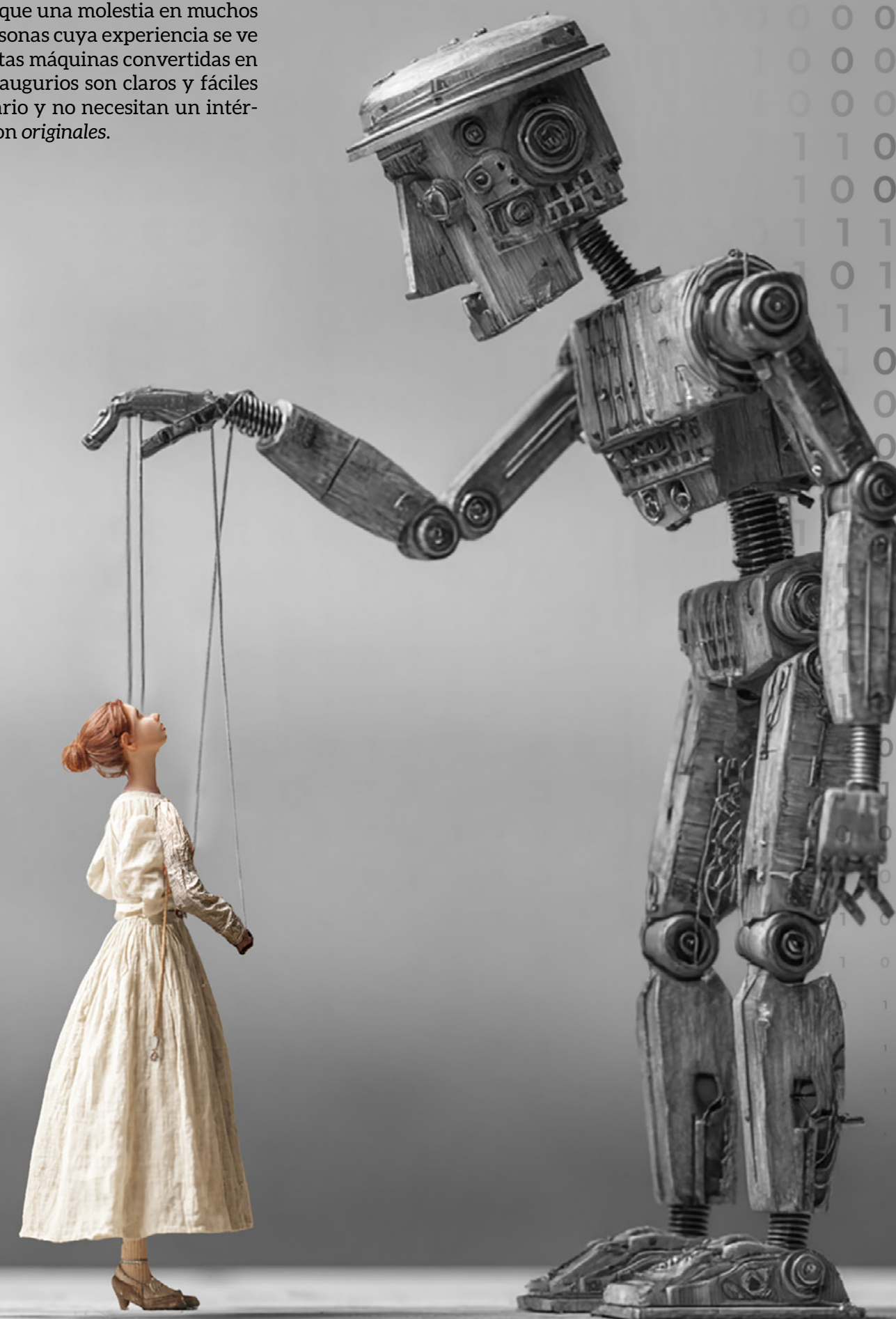
El uso extensivo de herramientas como ChatGPT o DeepSeek diseñadas para *generar* textos sofisticados indistinguibles de los que podrían ser producidos por un humano, son aplicables en una amplia gama de situaciones sociales y profesionales. Estas tecnologías presentan oportunidades, pero también, a menudo, desafíos éticos, legales, socioeconómicos y culturales y tienen el potencial de tener impactos tanto positivos como negativos para las organizaciones, la sociedad y las personas. Hoy es evidente que los modelos masivos del lenguaje han sacudido los cimientos de muchos ámbitos laborales.

Su funcionalidad básica de *pregúntame lo que quieras* se ha transformado en la creencia de que la máquina *siempre* tendrá una *buena* respuesta, *correcta* y *rápida*. Esta convicción sobre la capacidad de estas herramientas se ha convertido en algo más que una molestia en muchos ámbitos y para muchas personas cuya experiencia se ve desplazada por el uso de estas máquinas convertidas en oráculos de bolsillo, cuyos augurios son claros y fáciles de comprender por el usuario y no necesitan un intérprete, aunque casi nunca son *originales*.

Una publicación típica en redes sociales hablando de estas herramientas, suponiendo que no fuese una nota interesada, podría decir algo así:

Le pedí a ChatGPT/Llama que me escribiera un [texto/resumen/soliloquio/...] sobre [tema]. Y, a continuación, podríamos leer, DeepSeek hizo un trabajo [bueno/excelente/increíble], (sin duda) (mucho) mejor de lo que yo lo habría hecho.

Si bien muchas de esas anécdotas que circulan se relacionan con tareas *sencillas* de generación de contenido como la de corrector gramatical o traductor, estas publicaciones reflejan una clara tendencia a antropomorfizar los modelos de lenguaje: se les atribuyen cualidades humanas y capacidades creativas, como si realmente *hicieran* el trabajo por iniciativa propia, cuando en realidad son herramientas que operan a partir de instrucciones y datos proporcionados por los usuarios. Una de sus mayores ventajas es que estas respuestas, por complicadas que sean las preguntas, son gratis. Sin embargo, los resultados de estas máquinas deben verificarse cuidadosamente, ya que generan contenido inexacto, sesgado basado en fuentes de ideas incorrectamente reportadas, etc.



El texto del epígrafe de esta nota fue producido por una de estas herramientas, a mí me parece carente de emoción -un hallazgo *feliz* logrado no sabemos cuántos intentos después- y cuando lo comparo con una definición que el poeta mexicano José Gorostiza hace de la inteligencia:

*¡oh inteligencia, soledad en llamas!/
que lo consume todo hasta el silencio,/
sí, como una semilla enamorada/
que pudiera soñarse germinando!*

Expresión que está llena de esa íntima experiencia que produce el pensamiento en los humanos cuando están ensimismados y en intensa actividad creativa y que expresa esa emoción introspectiva, metafísica, que llega al silencio y, entonces, la semilla-idea se ve ya flor. Aunque una máquina pudiese escribir las mismas palabras, en el mismo orden jamás podría sentir esa emoción única pero que reconocemos por haberla experimentado.

En este contexto, las ideas del filósofo Luciano Floridi enriquecen y actualizan la analogía de Butler. Floridi sostiene que el verdadero desafío no es la tecnología en sí, sino la gestión y el control de la sociedad digital, misma que está en manos de unas pocas empresas, y los gobiernos han perdido la iniciativa dejando a los ciudadanos al arbitrio de unos pocos señores feudales digitales⁵.

Sin embargo, las máquinas nos ganan terreno; día a día estamos más subordinados a ellas; cada vez más personas se ven obligados a servirlos con nuevas formas de esclavitud; cada día una inmensa mayoría de personas, en todo el mundo, dedican las energías de toda su vida al desarrollo de la vida *mecánica o digital*.

Para J. Lanier⁶, el problema central reside en la manera en que estas tecnologías han sido diseñadas para recolectar información y manipular comportamientos, a menudo sin transparencia ni responsabilidad democrática. Los modelos masivos del lenguaje, como los mencionados en el texto, son tecnologías de amplio espectro y muy potentes, resulta imposible prever los posibles usos que los humanos, cualquier persona, pueda hacer y menos aún prevenir sus consecuencias ni atajarlas. Esto resulta más alarmante si, además, el usuario en cuestión carece de sentido crítico y cree a pie juntillas cualquier resultado que la herramienta produzca.

*¡Oh inteligencia, soledad en llamas!
que lo consume todo hasta el silencio,
sí, como una semilla enamorada
que pudiera soñarse germinando.*

José Gorostiza

Entre las amenazas que plantean ChatGPT y otros bots similares se incluyen que son algoritmos de caja negra u opacos, que las respuestas que generan promueven la discriminación y sesgos, vulgaridad; que sus respuestas pueden contener contenido textual incorrecto, inventado y contenido multimedia falso. Además, que al producir respuestas pueden cometer infracciones contra los derechos de autor, plagio, etcétera.

Por lo tanto, es esencial que los gobiernos, organizaciones e individuos comprendan, gestionen y mitiguen los riesgos derivados de la adopción de sistemas basados en la IA de esta potencia. Si no actuamos con pensamiento crítico resultado es cuestión de tiempo en el que quienes posean esas máquinas y sus algoritmos ostentaran la verdadera supremacía tecnológica y económica sobre el mundo y sus habitantes, es algo que ninguna persona con una mente filosófica puede cuestionar ni por un instante.

Es lícito, en 2025, preguntarse si no es demasiado tarde para detener la automatización y la digitalización. Sus defensores abogan que son aceleradores útiles para el desarrollo de la ciencia, habrá que verlo. Quizás este es el tiempo para que las sociedades se apliquen en pensar cómo reconducir el acceso a estas herramientas, hoy es libre y casi gratuito, antes de que alguien se haga daño o produzca males de forma *accidental*, gracia a una alucinación de una de estas máquinas.

ULISES CORTÉS

Catedrático de Inteligencia Artificial de la Universitat Politècnica de Catalunya. Coordinador Científico del grupo High-Performance Artificial Intelligence del Barcelona Supercomputing Center. Miembro del Observatori d'Ètica en Intel·ligència Artificial de Catalunya y del Comitè d'Ètica de la Universitat Politècnica de Catalunya. Es miembro del comité ejecutivo de EurAI. Participante como experto de México en el grupo de trabajo Data Governance de la Alianza Global para la Inteligencia Artificial (GPAI). Doctor Honoris Causa por la Universitat de Girona.



REFERENCIAS

- 1 <https://letralia.com/editorialletralia/especiales/bestiariaoartificial/2024/05/31/haiku-3-0/>
- 2 Thomas Hobbes, Leviathan, ed. C. B. Macpherson, London: Penguin, 1981, p. 150.
- 3 <https://mediarep.org/server/api/core/bitstreams/e0da505d-200c-43ab-be4b-6604a4df816f/content>
- 4 El entrenamiento desde cero es más costoso que la inferencia o el ajuste fino. Se requieren clústeres de GPU (a menudo cientos o miles para los modelos más grandes), con costos que alcanzan millones de euro para los LLM de última generación. Sin embargo, con la cuantificación, e.g. OmniQuant, y métodos eficientes de ajuste fino, se pueden reducir los requisitos de recursos para la personalización y la implementación.
- 5 <https://archive.org/details/technofeudalism-what-killed-capitalism-2023-yanis-varoufakis>
- 6 <https://observer.co.uk/style/features/article/you-me-and-chatgpt>

